EXAMINING THE PREVALENCE OF CRIMINAL DESISTANCE*

ROBERT BRAME
University of South Carolina

SHAWN D. BUSHWAY
RAYMOND PATERNOSTER
University of Maryland

Criminological theorists and criminal justice policy makers place a great deal of importance on the idea of desistance. In general terms, criminal desistance refers to a cessation of offending activity among those who have offended in the past. Some significant challenges await those who would estimate the relative size of the desisting population or attempt to identify factors that predict membership in that population. In this paper, we consider several different analytic frameworks that represent an array of plausible definitions. We then illustrate some of our ideas with an empirical example from the 1958 Philadelphia Birth Cohort Study.

KEYWORDS: Criminal careers, desistance, split population models.

INTRODUCTION

The National Research Council’s report on criminal careers identifies onset, frequency, seriousness, and duration as important dimensions of individual offending histories that merit careful scientific study (Blumstein, Cohen, Roth, and Visher, 1986). While a certain amount of controversy has surrounded the value of the criminal career framework (Gottfredson and Hirschi, 1990:240-241), detailed literatures have developed in the areas of participation and frequency in the years since the report was issued. This work has essentially focused on whether the processes that led to participation in or onset of offending are similar to the processes that generate variation in the frequency of offending. As a result of the interest in this area, empirical research in criminology has moved to the use of individual-level prospective data, and a number of methodological advancements such as so-called “trajectory models” and

* Note: The authors are members of the National Consortium on Violence Research.
growth curve models have allowed researchers to carefully study longitudinal patterns of offending (Osgood and Rowe, 1994; Nagin, 1999).

The literature on desistance, while considerably smaller than the literature on onset and frequency, has been motivated by a similar question: are the processes that lead to onset and frequency similar to the processes that lead to desistance? But there is a lack of consensus about the operational definition of the term "desistance" (Laub and Sampson, 2001). This lack of consensus is deeper than the usual debates about whether to use self-report or official records, or whether to include minor crimes in the definition of offending. Some researchers such as Farrington and Hawkins (1991) hope to study something approximating the permanent end of offending, while others, such as Elliot et al., (1989) and Clarke and Cornish (1985), view desistance as including temporary lulls in offending. This lack of consensus is driven in large part by the difficulty inherent in measuring desistance since it involves the absence of offending (Maruna, 2001), in combination with data collection efforts that have only followed individuals through adolescence and/or early adulthood. Longer follow-up periods would help resolve some of this ambiguity but the fact remains that researchers addressing the same question have invoked different conceptual and operational definitions of desistance.

We believe this lack of a consensus about what the field means by desistance has consequences for the development of a reliable understanding of desistance (Laub and Sampson, 2001). Bushway et al. (2003) demonstrate that the use of two different definitions on the same data set identify different sets of people as desisters. An empirical analysis of the causes of desistance could, therefore, lead to different conclusions depending on the definition of desistance used by the researcher. We believe that the development of a rich and meaningful empirical literature identifying the causes

---

2. The contemporary stream of literature on this topic dates back to the seminal work by Blumstein and Moitra (1980) and Blumstein et al. (1985). Since that time, a number of studies have focused explicitly on describing and/or modeling criminal desistance. A long but still incomplete list would include Mulvey and LaRosa (1986), Rand (1987), Fagan (1989), Paternoster (1989), Farrington and Hawkins (1991); Loeber et al., (1991); Shover and Thompson (1992); Elliott (1994); Sommers et al., (1994); Pezzin (1995); Shover (1996); Laub et al., (1998); Uggen and Kruittschnitt (1998); Uggen and Pilisvin (1998); Warr (1998); Paternoster et al., (2001); and Bushway et al. (2001). Recent detailed reviews of the issues raised by this body of research are presented in Laub and Sampson (2001) and Bushway et al. (2001).

3. The chance that someone changes in some important way after the follow-up period has terminated will always remain a problem in desistance studies with limited follow-up periods (Lehoczky, 1986:385; Laub and Sampson, 2001:9).

4. For detailed examples, see Laub and Sampson (2001:8), or Bushway et al. (2003).
of desistance requires a consensus about the key features of an operational definition of desistance.

Some important steps have been taken which should help move the field in this direction. Weitekamp and Kerner (1994) introduced the idea of separating the event of termination – the permanent end of criminal offending – from the process of desistance, during which the frequency of offending declines until termination. Moreover, there is a growing body of work, starting with Fagan (1989), that emphasizes the importance of studying desistance as a process of change that takes place over time. Laub et al. (1998) were the first to suggest that dynamic statistical models on prospective panel data can be used to examine the process of desistance. This approach places greater emphasis on modeling changes in offending over time and less emphasis on identifying the terminating event (Laub and Sampson, 2001:54).

Criminologists such as Land (1992), Nagin and Land (1993), and Osgood and Rowe (1994), among others, have been attempting to improve our ability to describe the processes of offending over time by introducing new descriptive statistical techniques for the study of offending over the life course. We believe these techniques, arising from the criminal career literature, can be productively employed to study desistance. But each of these models carries certain parametric assumptions about the nature of the desistance process. It is important, therefore, to ask not only whether the probabilistic modeling approach rising from the criminal career literature represents an improvement over a strict behavioral approach to desistance, but also whether some probabilistic models might be more appropriate than others.

In this paper, we apply several different statistical models to the criminal history records of the men in the 1958 Philadelphia Birth Cohort Study. To keep things simple, we set aside the important task of describing how desistance unfolds over time and emphasize instead the application of probabilistic models to the adult arrest distribution among those who experienced police contacts as juveniles. Specifically, we use several different models, including the standard behavioral model to estimate the proportion of the population who offended at least once before age 18 who can be described as "desisters" by age 27. Since each model in this paper uses a different operational definition of termination, each one produces a different estimate of the prevalence of termination. Such estimates are helpful to criminologists because they provide a foundation for the estimation of conditional proportions or probabilities that allow us to

5. For detailed discussions of the idea of desistance as a process see also Baskin and Sommers (1997); Bushway et al. (2001), Laub and Sampson (2001), and Maruna (2001).
reach a better understanding of the circumstances in offenders' lives that lead to diminished involvement in crime (Uggen and Piliavin, 1998:1407-1415; Laub and Sampson, 2001:54-55). If the estimates produced by the different models lead to substantively different conclusions about the prevalence of termination, criminologists must select the most suitable sets of operational definitions and assumptions upon which to base their conclusions. In this paper, we pay particular attention to the substantive meaning of these definitions and assumptions.6

**ANALYTIC FRAMEWORKS FOR STUDYING CRIMINAL DESISTANCE**

Several approaches to the quantitative study of desistance have occupied an important place in the criminological literature. These approaches generally assume that a population of individuals who have offended in the past is available for study. They also assume that the offending behavior of this population can be studied over the duration of some well-defined follow-up period. On the basis of the information provided by such data, the various approaches used in the criminological literature attempt to: (1) estimate the proportion of individuals who have desisted within the population of interest; and/or (2) study the statistical association between various life circumstances or background factors and the likelihood of desistance from crime. The primary difference between most studies of desistance lies in how they accomplish the first task. We will divide the methodologies into three categories: (1) strict behavioral desistance (individuals are categorized as desisters or persisters depending on whether they offend during the follow-up period); (2) approximate desistance (individuals are categorized as desisting if their offending frequency falls to a relatively low level, but possibly not exactly zero); (3) and split-population desistance (like (1) above, some individuals terminate while others do not but, unlike (1), the model takes into account the uncertainty associated with studying termination during a finite follow-up period.)

---

6. Laub and Sampson (2001:9) indicate that many difficulties confronting the measurement of criminal desistance also confront the study of individual involvement in criminal activity more generally. This raises important and difficult issues about whether criminal desistance should be studied with official records or self-reports, the appropriate range of behaviors that fit within the concept of "criminal behavior," and whether similar behaviors are interpreted consistently by different social and demographic groups (see e.g., Farrington and West 1995; Nagin et al., 1995; Uggen and Kruttschnitt 1998; Sommers et al., 1994; Warr 1998; Nielsen 1999; Benson 2001). Our objective in this paper is not to resolve these issues but rather to explore the implications of studying desistance through different analytical windows while assuming that the other difficulties in the study of criminal desistance are non-problematic. This approach allows us to focus our attention on variation in our results under different analytic frameworks while holding measurement issues constant.
Under the strict behavioral desistance approach, researchers create an arbitrary cut off point, usually an age, say 18, which is used to identify “offenders”. The selection of the cutoff point varies from dataset to dataset. Researchers then create a binary or dichotomous outcome variable that is coded 1 if the individual “persists” and 0 if the individual “terminates” within some time frame after the cutoff period, ranging from 1 year to 11 years in the current literature (e.g., Shover and Thompson, 1991; Smith and Brame, 1994; Warr, 1998; Farrington and Hawkins, 1991; Elliot et al., 1989; Loeber et al., 1991). These analyses typically then identify factors that can predict desistance within the time frame chosen in that particular study. However, their ability to identify the processes that produce an enduring criminal desistance is limited. There are at least two important shortcomings. First, the strict behavioral desistance approach does not allow for the possibility that some individuals who do not offend during a fixed period of time have not actually changed their behavior. This, of course, is the well-documented “false desistance” problem described by Greenberg (1991:18-19) and Laub and Sampson (2001:9). Second, the strict behavioral desistance approach does not do a good job of describing patterns of offending over time. As a result, this approach leads to difficulty in accomplishing the second task of linking the causes of desistance to the process of desistance (Bushway et al., 2003).

In contrast to the strict behavioral desistance approach, researchers have devised methods for more fully describing the pattern of offending over time by incorporating additional information that is available in many criminological data sets on the timing or frequency of criminal events within a fixed period of time into statistical models of offending (Maltz, 1984; Blumstein et al., 1985; Schmidt and Witte, 1988; Blumstein et al., 1988a, 1988b; Rowe, Osgood, and Nicewander, 1990; Greenberg, 1991; Laub, Nagin, and Sampson, 1998; Paternoster et al., 2001; Bushway et al., 2003). Methods emphasizing timing typically rely on survival time models while methods emphasizing the frequency at which events occur typically rely on event count models. In each case, the fundamental parameter of interest is the rate at which events occur during the follow-up period.

These probabilistic models all share four common features identified by Osgood and Rowe (1994:550): “(1) a curvilinear function linking the scale of the linear model and the scale of the measure of offending, (2) a probabilistic relationship between the latent propensity and the measured outcome, (3) a distribution of the latent individual differences and (4) relationships among repeated observations for the same individual.” While these models are now ubiquitous in published studies of offending on prospective data, we fear that the substantive implications of these features are not well understood. The second point identified by Osgood and Rowe is perhaps the most important. Probabilistic models of desistance
attempt to explain variation in an unobserved latent variable, commonly understood by statisticians as the propensity to offend.\footnote{The idea of a latent propensity to offend as described above is a simple and straightforward concept. It assumes that humans are essentially probabilistic, rather than deterministic, actors. Readers interested in a more detailed discussion of this issue are referred to Bushway et al. (2001) and Laub and Sampson (2001).} In the simple Poisson model of offending frequency, this propensity to offend is parameterized as the rate of offending, lambda ($\lambda$). We do not observe this parameter — rather we observe offending behavior. This parameter captures the systematic (non-random) forces that contribute to offending in any period. Yet, not all individuals with the same underlying propensity to offend will have the same observed level of offending because there are other non-systematic (probabilistic) forces at work, including random arrivals of opportunities and chance. As a result, we need a way to translate the underlying propensity to offend into observed behavior. This translation is the role of various probabilistic functions such as the Poisson distribution. Each probabilistic law makes certain known assumptions about the distribution of the underlying propensity to offend in the population (point three) and how offending in any given period affects offending in the next period (point four).

Points three and four represent the backbone of much of current criminological research. Point three refers to stable individual differences in offending over time. These differences lead to selection effects that always challenge causal analyses of criminal desistance based on longitudinal data (Uggen and Piliavin, 1998). Point four refers to the concept of state dependence (Nagin and Paternoster, 1991, 2000), which describes the ways in which offending in one period can lead to offending in the next period. Many developmental criminology theories (e.g., Thornberry, 1987; Sampson and Laub, 1993) are theories of state dependence — they describe the process by which current offending can have a causal effect on future offending.

Similarly, models of offending also make assumptions about the possibility of termination. Standard survival time and event count models like those cited above do not allow the offending rate to vanish (Schmidt and Witte, 1988; Greenberg, 1991). In other words, these models will allow the offense rate to approach but not reach zero (hence, the idea of approximate desistance) (Bushway et al., 2001). This is not a trivial matter because a non-vanishing offense rate contradicts the idea of desistance as a permanent cessation of offending activity (Blumstein et al., 1988a, 1988b; Greenberg, 1991:33-36).

It is possible to elaborate on the standard survival time and event count models...
modeling frameworks to allow for a mixture of two groups within a population of individuals who have offended in the past (i.e., a split-population model). One group—the desisters—is characterized by an offense rate of exactly zero while the other group—the persisters—is characterized by some non-zero offense rate that is allowed to approach but not reach zero (Schmidt and Witte, 1988:66-69; Greenberg, 1991:36). This elaboration can be implemented with both survival time models and event count models and it recognizes the fact that the follow-up period is finite in length.8 This so-called “split-population” model is a priori the most appealing framework for the study of desistance in the present context because we can use the timing and event count information available in the data to directly estimate the proportion of individuals who have terminated, while still describing the overall pattern of offending.

All of these modeling strategies have parameterizations that are consistent with the idea that there is a group of “desisters” in either a literal or an approximate sense. We are interested in investigating whether an estimate of the size of this group depends in important ways on the choice of particular definitional schema, or, alternatively, whether a variety of plausible strategies all lead to similar inferences. We will also examine whether the use of three different functional forms—Poisson, geometric and negative binomial—lead to different conclusions about the processes underlying the offending rate. Specifically, we will apply several plausible statistical models to data from the 1958 Philadelphia Birth Cohort Study (Tracy et al., 1990). To the extent that any particular model fits the data better, we have evidence to suggest that the assumptions of that model are a better reflection of the processes that drive offending (and therefore the desistance process).

DATA AND RESEARCH DESIGN

The 1958 Philadelphia birth cohort data are widely known to the criminology community and are among the best available prospective data for examining long-term officially recorded involvement in criminal behavior.

8. With survival time frameworks, the elaboration results in what Maltz (1984) calls the “incomplete failure time model” while the same elaboration has been called the “split-population failure time model” by Schmidt and Witte (1988). With event count frameworks, the elaboration results in what Lambert (1992) calls the “zero-inflation model” while Mullahy (1986, 1997) calls it the “with-zeroes model” and Greene (1997:945) calls it a “split-population model.” In both frameworks, the observed rate of event occurrence is assumed to be generated by two groups of people: (1) those whose probability of committing a crime in the follow-up period is exactly zero; and (2) those whose probability of committing a crime in the follow-up period is non-zero. For the second group we estimate the rate of event occurrence.
Details regarding the original study and data collection efforts are discussed in Tracy et al., (1990). For our analysis, we rely on data measuring both juvenile and adult police contacts for index offenses among the 13,160 males in the study. Because official records do not capture all criminal acts, it is reasonable to ask whether we can learn much about actual offending and desistance by using official records instead of self-reported offending. We believe that what Bushway et al., (2001) call "official" desistance – an end to involvement with the criminal justice system – is an interesting policy question separate from the question of absolute desistance. In the present context, policy makers might be very interested in knowing what causes individuals with a juvenile record to avoid contact with the adult criminal justice system once they reach age 18. But in order to answer that question, we need to develop a reasonable measure of official desistance. While the methods used here can be applied to self-report data, we think that applications to official data are likely to produce interesting and useful results as well.

Out of the 13,160 males in the initial cohort, a total of 2,657, or 20.2%, experienced at least one police contact prior to age 18. The majority of these individuals were nonwhite (73.8% nonwhite; 26.2% white) and below the median socioeconomic status score for the birth cohort (68.4% below median socioeconomic level; 31.6% above median socioeconomic level). In this paper, we examine the frequency distribution of adult police contacts between the ages of 18 and 27 for these 2,657 youths who offended as juveniles. Our main objective is to estimate the prevalence of desistance using different analytic models that include parameters with important conceptual linkages to the termination or cessation of offending activities. Table 1 (column #1) presents the frequency distribution of adult offenses among the 2,657 juvenile male index offenders from the Philadelphia Birth Cohort Study. Like most crime frequency distributions, this distribution is positively skewed with the majority of individuals (61.2%) in the data set exhibiting no further offending during the adult years (up to age 27). With these results in mind, we turn our attention to estimating the proportion of the Philadelphia juvenile offenders who terminate during adulthood.

9. It would be desirable to control for the effects of incapacitation or time "off the street" when studying this frequency distribution. Unfortunately, this information is not available in the Philadelphia data. The methods used in this paper, however, can be adapted to adjust for variation in street time (see e.g., King 1989:50, 124-126) and they could then be applied to data sets that do contain this information. We also caution readers that these desistance estimates should not be construed as typical of what we might expect to see in other data sets.
Table 1. Profile of Low-Rate Groups and Goodness-of-Fit Assessment (N = 2,657)

<table>
<thead>
<tr>
<th># of Adult Contacts</th>
<th>(1) Observed Distribution</th>
<th>(2) Poisson Rate Group</th>
<th>(3) Geometric Rate Group</th>
<th>(4) Three-Group Poisson Fit</th>
<th>(5) Two-Group Poisson Fit</th>
<th>(6) Split-Population Poisson Fit</th>
<th>(7) Split-Population Geometric Fit</th>
<th>(8) Negative Binomial Fit</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>.612</td>
<td>.833</td>
<td>.910</td>
<td>.617</td>
<td>.6123</td>
<td>.6123</td>
<td>.6123</td>
<td>.6099</td>
</tr>
<tr>
<td>1</td>
<td>.158</td>
<td>.153</td>
<td>.082</td>
<td>.1601</td>
<td>.1582</td>
<td>.1583</td>
<td>.1439</td>
<td>.1657</td>
</tr>
<tr>
<td>2</td>
<td>.082</td>
<td>.014</td>
<td>.007</td>
<td>.0759</td>
<td>.0816</td>
<td>.0809</td>
<td>.0905</td>
<td>.0838</td>
</tr>
<tr>
<td>3</td>
<td>.049</td>
<td>.001</td>
<td>.001</td>
<td>.0570</td>
<td>.0514</td>
<td>.0525</td>
<td>.0569</td>
<td>.0489</td>
</tr>
<tr>
<td>4</td>
<td>.040</td>
<td>------</td>
<td>------</td>
<td>.0370</td>
<td>.0334</td>
<td>.0357</td>
<td>.0358</td>
<td>.0305</td>
</tr>
<tr>
<td>5</td>
<td>.019</td>
<td>------</td>
<td>------</td>
<td>.0207</td>
<td>.0218</td>
<td>.0219</td>
<td>.0225</td>
<td>.0197</td>
</tr>
<tr>
<td>6</td>
<td>.013</td>
<td>------</td>
<td>------</td>
<td>.0113</td>
<td>.0143</td>
<td>.0125</td>
<td>.0141</td>
<td>.0130</td>
</tr>
<tr>
<td>7</td>
<td>.008</td>
<td>------</td>
<td>------</td>
<td>.0071</td>
<td>.0093</td>
<td>.0074</td>
<td>.0089</td>
<td>.0088</td>
</tr>
<tr>
<td>8</td>
<td>.003</td>
<td>------</td>
<td>------</td>
<td>.0053</td>
<td>.0061</td>
<td>.0051</td>
<td>.0056</td>
<td>.0060</td>
</tr>
<tr>
<td>9</td>
<td>.005</td>
<td>------</td>
<td>------</td>
<td>.0043</td>
<td>.0040</td>
<td>.0039</td>
<td>.0035</td>
<td>.0041</td>
</tr>
<tr>
<td>10</td>
<td>.002</td>
<td>------</td>
<td>------</td>
<td>.0033</td>
<td>.0026</td>
<td>.0031</td>
<td>.0022</td>
<td>.0029</td>
</tr>
<tr>
<td>11</td>
<td>.003</td>
<td>------</td>
<td>------</td>
<td>.0024</td>
<td>.0017</td>
<td>.0023</td>
<td>.0014</td>
<td>.0020</td>
</tr>
<tr>
<td>12</td>
<td>.002</td>
<td>------</td>
<td>------</td>
<td>.0016</td>
<td>.0011</td>
<td>.0017</td>
<td>.0009</td>
<td>.0014</td>
</tr>
<tr>
<td>13</td>
<td>.000</td>
<td>------</td>
<td>------</td>
<td>.0010</td>
<td>.0007</td>
<td>.0011</td>
<td>.0006</td>
<td>.0010</td>
</tr>
<tr>
<td>14</td>
<td>.001</td>
<td>------</td>
<td>------</td>
<td>.0006</td>
<td>.0005</td>
<td>.0007</td>
<td>.0004</td>
<td>.0007</td>
</tr>
<tr>
<td>15</td>
<td>.001</td>
<td>------</td>
<td>------</td>
<td>.0003</td>
<td>.0003</td>
<td>.0004</td>
<td>.0002</td>
<td>.0005</td>
</tr>
</tbody>
</table>

\[ \hat{T} \]

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Estimated (k)</th>
<th>5</th>
<th>3</th>
<th>6</th>
<th>2</th>
<th>2</th>
</tr>
</thead>
</table>

\[ p(\hat{T} > \chi^2 | T = 0, df = c - k - 1) \]

|                  | .142       | .121     | .247     | .003     | .111 |

Note: Expected frequencies below underlined cells are combined so that there is at least five expected cases per cell.
In this section, we present a number of different estimates of the proportion of individuals who desist in the Philadelphia cohort. We begin by considering the estimate implied by a strict behavioral model of desistance. Then, we consider two approximate desistance models that both provide a very good fit to the event count data in the first column of Table 1. Next, we consider two split-population models that explicitly allow for termination. We then consider an interesting possibility—a specification that formally rejects the idea that there is a discrete group of true desisters or even approximate desisters. Finally, we compare the performance of the different models.

**Strict Behavioral Desistance**

Based on the information in Table 1, one potential estimate of the prevalence of desistance would be 61.2%. This approach does not allow for a very detailed description of the process of desistance, and will most likely not lead to interesting causal analysis of the process of desistance. An absence of offenses among individuals who have offended in the past does not logically imply that the propensity to commit crimes has changed (Bushway et al., 2001:495; Greenberg 1991:19; Barnett et al., 1987, 1989). Indeed, the approximate desistance models discussed above take as axiomatic the idea that: (1) there is a nonzero probability of exactly zero offenses within a finite time period for any given rate of offending; and (2) the rate of offending must always be greater than zero. Because criminologists have been aware of this issue for some time (see e.g., Barnett and Lofaso, 1985; Schmidt and Witte, 1988; Rowe et al., 1990; Blumstein et al., 1988a, 1988b; Barnett et al., 1987, 1989), a number of models that fit within the “approximate desistance” and “split-population” analytic frameworks have been proposed to address it.

**Approximate Criminal Desistance**

In this section, we consider two models which allow for approximate criminal desistance. The first model is a mixture of three Poisson processes while the second model is a mixture of two geometric processes. We describe the key features of each model and the estimation results are presented in the top panel of Table 2.

**Mixture of Three Poisson Processes**

The Poisson mixture model has received wide use within the field of criminology over the past thirty years. A Poisson process is based on a set
Table 2. Desistance Models ($N = 2,657$)

### Approximate Desistance Specifications

<table>
<thead>
<tr>
<th>Three-Group Poisson Mixture</th>
<th>Two-Group Geometric Mixture</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated Contact Rate ($\lambda$)</td>
<td>Percent of Population ($\pi \times 100$)</td>
</tr>
<tr>
<td>Low-Rate Group</td>
<td>0.182</td>
</tr>
<tr>
<td>Medium-Rate Group</td>
<td>2.540</td>
</tr>
<tr>
<td>High-Rate Group</td>
<td>8.142</td>
</tr>
<tr>
<td>Log-likelihood = -3597.02</td>
<td>Log-likelihood = -3599.16</td>
</tr>
</tbody>
</table>

### Split-Population Desistance Specifications

<table>
<thead>
<tr>
<th>Split-Population Poisson Mixture</th>
<th>Split-Population Geometric Mixture</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated Contact Rate ($\lambda$)</td>
<td>Percent of Population ($\pi \times 100$)</td>
</tr>
<tr>
<td>Desisting Group</td>
<td>------</td>
</tr>
<tr>
<td>Low-Rate Group</td>
<td>0.537</td>
</tr>
<tr>
<td>Medium-Rate Group</td>
<td>2.945</td>
</tr>
<tr>
<td>High-Rate Group</td>
<td>8.611</td>
</tr>
<tr>
<td>Log-likelihood = -3595.34</td>
<td>Log-likelihood = -3603.91</td>
</tr>
</tbody>
</table>
of assumptions about human behavior including: (1) the rate at which events occur is constant throughout the population of interest; and (2) events arrive purely randomly in time at the Poisson rate. The first assumption implies that there is no population heterogeneity in the rate at which events occur and while the second assumption implies that there will be no time trend in the rate of offending and that past experience will have no influence on the rate in the future (i.e., no state dependence) (Nagin and Paternoster, 1991). Taken together the two assumptions imply that there is no termination of offending.

When events occur according to this process, we can make some concrete predictions about what the resulting event frequency distribution will look like so long as we know the rate which governs the process. The Poisson rate parameter, $\lambda$, can be viewed as the average frequency of events that occur within the follow-up period. With this estimate of the rate, we can calculate an expected frequency distribution under a Poisson process and compare it to the actual, observed frequency distribution. For crime count data, the Poisson process virtually never fits the observed data very well (see e.g., Blumstein et al., 1985; Blumstein et al., 1988a, 1988b; Barnett et al., 1987, 1989). Given the assumptions of the Poisson process, this lack of fit is not surprising. As suggested by Osgood and Rowe (1994), realistic models of criminal behavior will have to confront both population heterogeneity and state dependence.

To address this problem, statisticians have devised a variety of ways to relax some of the restrictive assumptions underlying the Poisson model. One of the simplest ways to do this is to assume that there are stable individual differences and divide the population of interest into several distinct groups of individuals, each of which has its own Poisson rate parameter. Our preliminary analysis indicates that a three group model provides the best fit to the observed data. Table 2 further indicates that the three groups identified in our Poisson mixture model analysis cover a considerable range of variation in adult contact frequency. The low rate group has an average of only about 0.18 contacts over the course of the ten years.

10. To estimate this model, we use the method of maximum likelihood with a Newton-Raphson optimization method. The likelihood function is given by:

$$L(y_1, y_2, \ldots, y_N | \pi, \lambda) = \prod_{i=1}^{N} \left( \sum_{j=1}^{3} \pi_j \frac{\exp(-\lambda_j) \times \lambda_j^{y_i}}{y_i !} \right)$$

where $y_i$ is the observed number of adult contacts for each of the $i = 1, 2, \ldots, N = 2,657$ individuals in the study, $\lambda_j$ is the rate at which events occur for individuals in group $j$ and $\lambda_j = \exp(y_j)$. The model also provides us with an estimate of the probability of group membership, $\pi$, for each of the three groups. The three group model fits better than the two group model when we compare the expected frequencies under the model to the observed frequency distribution.
year follow-up period and this group is estimated to include about 71% of the population. It is worth noting that our estimate of the prevalence of this low-rate group is not equal to the estimate obtained under our strict behavioral desistance definition. We encounter this dissonance because the model allows for the realistic possibility that some individuals who are in the mid-rate and high-rate groups will still have zero offenses because of the finite length of the follow-up period.

One potential drawback of the model is that it makes no distinction between individuals who offend at very low rates and individuals who have an absolute zero offense rate level (if such individuals exist). While the split-population models we consider later avoid this problem, it may turn out that there is nothing lost from thinking about low-rate offenders as a homogeneous group. To investigate the behavior of this low-rate group, column #2 of Table 1 also shows the expected frequency distribution of adult offenses for the low-rate group revealing that the label of "low-rate" is an appropriate one.

**Mixture of Two Geometric Processes**

Another useful model for investigating crime frequency distributions is based on the geometric probability distribution (Evans et al., 1993:82; Grandell, 1997:4). The geometric distribution assumes that each individual in the population can be characterized by the same probability of committing a first offense. Now, among those individuals that cross the hurdle of committing the first offense, they have the same probability of committing a second offense and so on. Within this simple framework, there is no concept of termination.

As with the above Poisson model, we can loosen the constraint of a constant probability of offending across the population by assuming that there are multiple groups. We found that the best-fitting mixture model for the geometric distribution was the two-group mixture model. As

11. The geometric and Poisson models are actually quite different specifications of the process by which event counts are generated. In fact, it is possible to derive a simple geometric model by allowing the Poisson parameter, $\lambda$, to vary according to an exponential probability distribution. It follows that the constraint of a constant offense rate, $\lambda$, is not equivalent to the constraint of a constant probability of committing the next offense.

12. We attempted to estimate both two and three group geometric mixtures but no improvement in fit resulted from the more complicated three group mixture. The likelihood function for the two group model is:

$$L(Y, \pi, \theta, y) = \prod_{i=1}^{N} \left( \pi_i \times \left( \theta_i \times (1-\theta_i)^y \right) \right) + \left( 1 - \pi_i \right) \times \left( \theta_i \times (1-\theta_i)^y \right)$$

where $y_i$ is the observed number of adult contacts for each of the $i = 1, 2, \ldots, N = 2,657$. $\theta_i$ is the probability parameter for individuals who are members of group $j$ and $\theta_i = \Phi(y_i)$
Table 2 indicates, this model results in the identification of a low-rate group with an average rate of 0.099 contacts, and a high-rate group with an average rate of 1.889 contacts. According to this model, the probability that an individual is a member of the low-rate "approximate desister" group is 0.472 which is substantially lower than the estimate obtained under the strict behavioral desistance definition. The third column of Table 1 also shows the expected frequency distribution of adult offenses for the low-rate group and it appears that this distribution is even more skewed than the three-group mixed Poisson distribution.

**SPLIT-POPULATION MODELS**

We now attempt to relax the "no termination" assumption by considering models that allow for the possibility that the population is comprised of two types of individuals: (1) a group of individuals who have exactly zero chance of committing any adult offenses (i.e., true desisters); and (2) a group of individuals whose crime frequency distribution is generated by a Poisson process governed by the rate parameter, \( \lambda \), discussed above. Everyone in the first group will have exactly zero offenses because they have zero probability of committing any crimes. These features of the model make it an attractive platform for formalizing ideas about termination (Schmidt and Witte, 1988:67-68; Greenberg, 1991:36).

Within the second group, some individuals will exhibit zero offenses but this outcome will only occur because of the natural variation that is created by a Poisson process. Individuals within this group who exhibit zero offenses, then, might properly be characterized as "false desisters," because if we only observed them long enough, we would have the opportunity to observe their "true colors" (Greenberg, 1991:22; Bushway et al., 2001:498-499).

We will consider two specific forms of these so-called split-population event count models in detail. The first model – a split-population Poisson mixture – allows for the existence of four groups of individuals including a group of desisters and three groups of individuals who have not desisted but have varying rates of involvement in adult criminal offending. The second model – a split-population geometric model – allows for the existence of two groups of individuals including a group of desisters and a group of individuals who continue to offend according to a geometric distribution.

where \( \Phi(\cdot) \) is the standard normal cumulative distribution function to ensure that \( \theta \) falls in the interval (0,1). The contact rate for individuals in group \( j \) is calculated by \( (1-\theta_j)/\theta_j \) (Evans et al., 1993:82). The model also provides us with an estimate of the probability of group membership, \( \pi_1 \) and \( \pi_2 = 1 - \pi_1 \), for each of the two groups.
EXAMINING CRIMINAL DESISTANCE

SPLIT-POPULATION POISSON MIXTURE MODEL

We estimated a number of different zero-modified Poisson mixtures including specifications that allowed for a group of desisters and varying number of groups whose offense frequency distributions are assumed to vary according to a Poisson process. The best of these models was one which allowed for a group of desisters and three groups of individuals with different offense rates. The results of the analysis are presented in the bottom panel of Table 2 and they reveal that about 36.6% of the population can be characterized as true desisters. This estimate is significantly lower than the proportion of people who actually have zero offenses because the individuals who have zero offenses are assumed to be a mixture of two groups: (1) those who have terminated; and (2) those who have not terminated but were simply not yet observed to have any offenses during the follow-up period.

SPLIT-POPULATION GEOMETRIC MODEL

The bottom panel of Table 2 also presents the results of our analysis using the geometric split-population model. This model assumes that there are two types of individuals: (1) people who have terminated and

13. The likelihood function for this model has two distinct parts. For individuals who have exactly zero offenses during adulthood the likelihood is:

\[ L(y_i, \delta, \pi, \xi) = \prod_{i=1}^{n} \Phi(\delta) + [1-\Phi(\delta)] \times \left( \sum_{j=1}^{k} \pi_j \times \exp(-\lambda_j) \right) \]

while the likelihood for those with one or more offenses during adulthood is:

\[ L(y_i, \delta, \pi, \xi) = \prod_{i=1}^{n} [1-\Phi(\delta)] \times \left( \sum_{j=1}^{k} \pi_j \times \frac{\exp(-\lambda_j) \times \lambda_j^{y_i}}{y_i!} \right) \]

where \( \delta \) is the so-called "splitting" parameter.

14. The probability that an individual is a true desister in this model is given by \( \Phi(\delta) \). On the other hand, the probability that an individual has zero contacts is estimated by:

\[ p(y_i = 0 | \delta, \pi) = \Phi(\delta) + [1-\Phi(\delta)] \times \sum_{j=1}^{k} (\pi_j \times \exp(-\lambda_j)) \]

Under this framework, a zero contact outcome can arise for two reasons: (1) an individual is a true desister; and (2) an individual has zero offenses by chance alone.

15. Like the split-population Poisson specification above, the likelihood function for the split-population geometric model is written in two parts. For those individuals with zero contacts in the follow-up period, the likelihood is:

\[ L(y_i = 0 | \delta, \xi) = \prod_{i=1}^{n} \Phi(\delta) + [1-\Phi(\delta)] \times \theta \]

while, for those individuals with at least one contact in the follow-up period, the likelihood is:

\[ L(y_i > 0 | \delta, \xi) = \prod_{i=1}^{n} [1-\Phi(\delta)] \times [\theta \times (1-\theta)^{y_i}] \]
who, therefore, do not have any new contacts; and (2) individuals who have not terminated and whose contact frequency distribution is approximately geometric. The estimated proportion of people who have desisted under this model is 38.3% which is actually very close to the estimate provided by the split-population Poisson mixture model. Thus, the two split-population models we have considered (there may be other useful models we have not considered) are in basic agreement about the relative size of the desisting population.

NEGATIVE BINOMIAL MODEL

The negative binomial specification differs in some important ways from the other models we have examined. Although there are several ways to derive a negative binomial distribution, one approach is to assume that each individual draws a Poisson rate parameter, \( \lambda_i \), from a gamma probability distribution. Then, for each individual, crime frequencies occur as a result of a Poisson process conditioned on that individual's offending rate.

As researchers like Rowe, Osgood, and Nicewander (1990) and Greenberg (1991) have noted, continuous Poisson mixtures like the negative binomial distribution drop the assumption that there are discrete groups of individuals in favor of an alternative assumption that the propensity to commit crimes is continuously distributed. This assumption implies that the propensity to offend can approach values arbitrarily close to zero even though an exact value of zero (i.e., desistance) is never attainable. With the models proposed by these authors, an individual always maintains some non-zero probability of offending activity even though the probability may be extremely small. A conceptual difficulty with this method, therefore, is that we cannot actually identify the proportion of people who terminate because all individuals are assumed to maintain at least some probability – however small – of continuing to offend in the future. In fact, under the assumptions of this model a discrete group of desisters does not exist, rather all offenders' propensity to commit crime simply diminishes continuously (Greenberg 1991:19). This does not pose any particular problem for these researchers, however, because they question whether it is realistic to classify individuals as desisters in the first place. Theoretically, this view is consistent with the position of Gottfredson and Hirschi (1990: 240-241) who have argued that initiation, frequency, and duration of offending activity all share the same common cause and that it is not, therefore, necessary to study them separately. As they see it, there is little merit in the idea that there is a distinct and meaningful group of people that can be referred to as "desisters", and there is little point in developing theories to account for them or policies which assume they exist. In the present context, this skepticism seems plausible,
so we examined the performance of the negative binomial distribution in comparison to the various finite mixture distributions we have already examined. The model used here was described in detail by Greenberg (1991:23) and the estimation results are presented in Figure 1.16

Based upon the parameter estimates, the graph in Figure 1 displays the expected distribution of the contact rate under the negative binomial assumptions. This graph suggests that the expected distribution of λ is highly skewed just like the adult offense frequency distribution. As we argued above, the obvious problem with this model is that it does not easily lend itself to a discussion of who can be properly characterized as a desister. Certainly, anyone is free to select a cutting point in the distribution in order to categorize some previous offenders as desisters if their offending rate is "close to" zero. In fact, the approximate desistance models described above provide just such a cutting point. But, this is only because the functional form of the model assumes that the distribution of the offending rate is discrete rather than continuous. With the negative binomial model, where the crime rate distribution is continuous rather than discrete, there is no natural cutting point and "any cutting point is bound to appear somewhat arbitrary and difficult to justify" (Greenberg 1991:40).

In sum, the most serious substantive issue with this mathematical model is that it fails to provide us with a clear framework for thinking about a discrete group of desisters. The fact that this difficulty exists, however, does not mean that the model is wrong. In light of this fact, it becomes important for us to determine which of the models we have estimated best fits the data.

16. The likelihood function is:

\[ L(\alpha, \beta | y) = \prod_{i=1}^{n} \frac{\Gamma(y_i + \alpha) \times \left( \frac{1}{\beta} \right)^{y_i+\alpha}}{y_i! \times \Gamma(\alpha) \times \left( 1 + \frac{1}{\beta} \right)^{y_i+\alpha}} \]

where \( p(\lambda) \sim \Gamma(\alpha, \beta) \), and \( \Gamma(\cdot) \) is the gamma function (i.e., for integer values, \( \Gamma(x+1) = x! \)). As Greenberg (1991:23) shows, this parameterization of the negative binomial distribution produces a maximum likelihood estimate of the average value of λ for the population under study as well as its variance which are given by:

\[ \bar{\lambda} = \frac{\alpha}{\beta} \quad \text{and} \quad V(\lambda) = \frac{\alpha}{\beta^2}. \]
COMPARING THE MODELS

We now have several different estimates of the size of the desisting population from our analysis of the Philadelphia data. Since these models are all plausible specifications for the process that produces criminal desistance, a comparison is necessary. On the surface, Table 1 (columns 4-8) suggests that all of these specifications provide what might be called a "good" fit to the observed data. Nevertheless, upon close inspection of Table 1 some variation emerges. Based on the information in this table, the split-population Poisson model emerges as the best-fitting specification while the split-population geometric model appears to be the worst. Interestingly, as suggested in Table 2, these two models produced very different
estimates of the size of the desisting population. The three-group Poisson, two-group geometric, and negative binomial models occupy the middle ground. The split-population and two-group geometric models perform relatively poorly in the tail of the distribution while the negative binomial model performs relatively poorly in fitting the number of individuals with four adult offenses.

Because the different models have different numbers of parameters, it is also useful to consider penalizing the models based on the number of parameters estimated. One way to do this is to calculate a chi-squared statistic \( \chi^2 \) and refer it to a chi-squared table with \( c-k-I \) degrees of freedom, where \( k \) is the number of parameters estimated and \( c \) is the number of cells in the table. As the bottom row of Table 1 suggests, the split-population Poisson model still performs best and the split-population geometric model still performs worst under this comparison. The three-group Poisson, two-group geometric, and negative binomial models still occupy the middle ground. Based on this evidence, we reach the conclusion that the split-population Poisson model fits the data the best and in the absence of an alternative specification that fits better, we are inclined to put the most weight on the estimate of 36.6% desistance prevalence produced by that estimator. Nevertheless, our conclusion is limited by the understanding that all of the models fit the data quite well and that they produced very different inferences about the size of the desisting population.

17. We wish to thank an anonymous reviewer of a previous version of the manuscript who pointed out to us that the oft-used Bayesian Information Criterion (BIC) has recently become controversial (Winship 1999). When we originally submitted this paper for review we reported that the BIC was optimized by the negative binomial model. As our reviewer pointed out, however, the reason for this appears to be that the BIC extracts a penalty for the number of parameters estimated that increases substantially when the sample size becomes large. In the case of the Philadelphia data, the BIC places less weight on the lack of fit (particularly at four offenses) and greater weight on the fact that the fit (which was not a bad fit) was achieved with only two parameter estimates. Nevertheless, the reviewer suggested that we use the Akaike Information Criterion (AIC) instead. The AIC extracts a penalty for the number of parameters estimated but this penalty is not affected by the sample size. When we did this, we achieved results that are the same as those based on the goodness-of-fit comparison we reported in the bottom rows of Table 1. In light of the controversy surrounding the use of the BIC and the correspondence between our goodness-of-fit assessments and those based on the AIC, we prefer to rely on relatively simple goodness-of-fit measures to assess the performance of the different models in this paper.

18. One of the manuscript reviewers expressed concern about the large disparity between the desistance estimate provided by the strict behavioral model (61.2%) and the desistance estimate provided by the split-population Poisson model (36.6%), given that all individuals in the strict behavioral model have not offended for 9 years. The difference can be attributed to the need to differentiate between very low levels of offending and a true zero rate of offending under the assumptions of the zero-inflated
DISCUSSION AND CONCLUSIONS

In this section, we sought to investigate one area of difficulty confronting desistance researchers – the problem of rigorously defining the meaning of the term. We began this research with a simple question: would researchers obtain different results if they relied on different – and plausible – formal definitions of desistance? To focus on this issue, we proposed three different analytical frameworks for thinking rigorously about the meaning of desistance, all of which have been used in studies that relate to the issue of criminal desistance.

The simplest of these frameworks is the idea of strict behavioral desistance. Under this model, individuals who are observed to offend within a finite follow-up period are viewed as persisters while individuals who refrain from offending are viewed as desisters. Although this common approach has the dual advantages of simplicity and interpretability, it also suffers from the requirement that individuals be classified as desisters or persisters based on behavior over a limited period of time, which leaves it open to the problem of “false desistance.”

The other two frameworks we examined – the approximate desistance and split-population models – both accomplish the goal of studying desistance probabilistically. Each of our proposed statistical models is comprised of a formal probabilistic function which translates the latent propensity to offend into observed behavior. These functions carry with them assumptions about offending patterns in four key areas: (1) the existence of stable individual differences in the rate of offending; (2) the relationship between past and future offending behavior; (3) the existence of termination (“true desisters”); and (4) the extent to which the underlying propensity to offend varies over the observation period.

For example, our two models of “approximate desistance” are based on mixtures of Poisson and geometric distributions. These models assume that contact frequencies follow either a Poisson or geometric distribution after conditioning on whether one is a low-rate or high-rate offender. Although the simple versions of these models are unrealistic for the study of desistance, we rely on mixture models which allow for variation in the Poisson rate and the geometric probability parameters (Lehoczky 1986). With this modification, the Poisson and geometric models provided excellent fits to the observed adult contact frequency distributions in the 1958 Philadelphia birth cohort data. In addition, both mixture models provide plausible mathematical definitions of desistance. In both instances, the

Poison model. Lehoczky (1986) earlier speculated that such an effort has limited utility and recommended against it. On the other hand, the approach in this paper highlights our ability to match the definition of true desistance to a statistical model when we pay attention to the assumptions of the model.
models allow for a group of individuals who exhibit "very low" contact levels in adulthood. A reliance on this modeling framework, then, presupposes that a researcher is willing to abandon the requirement of a complete absence of offending as a standard for treating individuals as "desisters."

The final definition retains the idea that there is a group of individuals who terminate while accounting for the reality that we can typically observe human subjects over a finite period of time. Our Poisson and geometric split-population models accomplish this dual objective and, as Table 1 suggests, they also fit the observed data.

Despite the fact that each of the probabilistic models fit the data well, they all produced quite different estimates of the prevalence of termination. Our strict behavioral definition of desistance resulted in an estimated desistance probability of 61.2% among those who offend at least once before age 18. On the other hand, our approximate desistance model estimates varied widely depending on whether we accept the Poisson definition (prevalence rate: 71.0%) or the geometric definition (prevalence rate: 47.2%). Finally, our split-population models were much closer to agreement with each other (Poisson prevalence: 36.6%; geometric prevalence: 38.3%).

We draw several conclusions from this result. First, models and their assumptions matter. Each of the models had a number of assumptions about the existence and form of desistance, and they lead to very different inferences about the prevalence of desistance. The split population models' assumptions are most in line with the traditional definition of termination or true desistance and they provide the smallest estimate of the size of the desisting population. The other models, which have less conservative definitions of desistance, predictably provide larger estimates for the prevalence of desistance. Second, it is possible to identify "true desisters" in a manner consistent with the way Blumstein and other criminal career researchers have defined the term by using split-population models. Although split-population models have not been used frequently in the criminology literature, we believe this paper has demonstrated both the feasibility and attractiveness of this approach. Third, "false desistance," a commonly expressed concern with the behavioral approach to desistance, is an important issue. Relying only on observed behavior without taking individuals' underlying propensity to offend into account would produce significant overestimates of the existence of adult desistance. Based on the evidence presented in Table 1, we believe that the split-population Poisson model provides the best estimate of "true desistance" in the Philadelphia data set which we fix at 36.6%.

Of course, this analysis is confined to one data set based exclusively on official record information. The split-population Poisson model may not
be the best model for self-report data, or even with other official record
data sets. We think a productive line of future research would involve
estimation work in other data sets based on official record and self-report
data.

We also advocate an extension of this approach to the dynamic trajec-
tory models currently being used to study the process of desistance (Laub,
et al., 1998; Bushway, et al., 2003). Like our models of approximate and
ture desistance with the retrospective cross-sectional data used in this
paper, each trajectory model is based on different parametric assumptions
which could affect inferences about the processes that lead to desistance
from crime. We believe that the study of desistance will benefit from
efforts to explore the properties of dynamic models as well as the static
ones examined here.

REFERENCES

Barnett, Arnold., Alfred Blumstein, and David P. Farrington
Barnett, Arnold., Alfred Blumstein, and David P. Farrington
Barnett, Arnold and Anthony J. Lofaso
1985 Selective incapacitation and the Philadelphia cohort data. Journal of Quan-
titative Criminology 1:3–36.
Benson, Michael L.
Blumstein, Alfred, Jacqueline Cohen, Jeffrey Roth, and Christy A. Visher
Blumstein, Alfred, Jacqueline Cohen, and David P. Farrington
Blumstein, Alfred, Jacqueline Cohen, and Soumyo Moitra
Blumstein, Alfred, David P. Farrington, and Soumyo Moitra
1985 Delinquency careers: innocents, desisters, and persisters. In Michael Tonry
Blumstein, Alfred, and Soumyo Moitra
1980 The identification of “career criminals” from “chronic offenders” in a
Bushway, Shawn D., Alex R. Piquero, Lisa M. Broidy, Elizabeth Cauffman, and Paul
Mazerolle

Bushway, Shawn D., Terrence P. Thornberry, and Marvin D. Krohn

Elliott, Delbert S., David Huizinga, and Scott Menard
Multiple Problem Youth: Delinquency, Substance Use and Mental Health. New York: Springer-Verlag.

Elliott, Delbert S.
Serious violent offenders: onset, developmental course, and termination. Criminology 32:1–22.

Evans, Merran, Nicholas Hastings, and Brian Peacock

Fagan, Jeffrey

Farrington, David P. and J. David Hawkins
Predicting participation, early onset, and later persistence in officially recorded offending. Criminal Behavior and Mental Health 1:1–33.

Gottfredson, Michael R. and Travis Hirschi

Grandell, Jan

Greene, William H.

Greenberg, David F.

King, Gary

Lambert, Diane

Laub, John H., Daniel S. Nagin, and Robert J. Sampson

Laub, John H. and Robert J. Sampson

LeBlanc, Marc and Rolf Loeber
Lehoczky, John P.

Loeber, Rolf and Marc LeBlanc

Loeber, Rolf, Magda Stouthamer-Loeber, W. Van Kammen, and David P. Farrington

Maltz, Michael D.

Mullahy, John

Mulvey, Edward P. and John F. LaRosa

Nagin, Daniel S.

Nagin, Daniel S. and Kenneth Land

Nagin, Daniel S., David P. Farrington, and Terrie E. Moffitt
1995 Life-course trajectories of different types of offenders. Criminology 33:111–139.

Nagin, Daniel S. and Raymond Paternoster

Nagin, Daniel S. and Douglas A. Smith

Nielsen, Amie L.

Paternoster, Raymond
Paternoster, Raymond and Ruth Triplett  

Paternoster, Raymond, Robert Brame, and David P. Farrington  

Pezzin, Liliana E.  

Rand, Alicia  

Rowe, David C., D. Wayne Osgood, and Alan W. Nicewander  

Schmidt, Peter and Ann D. Witte  

Shover, Neal  

Shover, Neal and Carol Y. Thompson  

Sommers, Ira, Deborah R. Baskin, and Jeffrey Fagan  

Smith, Douglas A. and Robert Brame  

Tracy, Paul, Marvin Wolfgang, and Robert M. Figlio  

Uggen, Christopher and Candace Kruttschnitt  

Uggen, Christopher and Irving Piliavin  

Warr, Mark  

Weitkamp, Elmar G.M. and Hans-Jurgen Kerner  
Winship, Christopher
1999 Editor's introduction to the special issue on the Bayesian Information Criterion. Sociological Methods and Research 27:355-358.

Robert Brame is an associate professor in the Department of Criminology and Criminal Justice at the University of South Carolina. His research interests center on juvenile delinquency and criminological theory.

Shawn D. Bushway is an Assistant Professor of Criminology in the Department of Criminology and Criminal Justice at the University of Maryland. He received his Ph.D. in Public Policy Analysis and Political Economy in 1996 from the H. John Heinz III School of Public Policy and Management at Carnegie Mellon University. His current research focuses on understanding the process of desistance and disparity in sentencing outcomes.

Raymond Paternoster is Professor in the Department of Criminology and Criminal Justice at the University of Maryland. His research interests include criminological theory, quantitative methods in criminology, and issues related to the death penalty.